

Patent number: JP7273922
Publication date: 1995-10-20
Inventor: ITO SATORU; others: 02
Applicant: MATSUSHITA ELECTRIC IND CO LTD
Classification:
- international: H04N1/00; G06K9/00
- european:
Application number: JP19940060762 19940330
Priority number(s):

[View INPADOC patent family](#)

Abstract of JP7273922

PURPOSE:To implement transfer in matching with the importance by checking the contents freely written on a facsimile transmission table for each item.

CONSTITUTION:A character recognition means 2 converts a facsimile text reached to a reception section 1 into a text and a keyword extract means 3 extracts a keyword to improve the character recognition rate and to summarize the content. An item discrimination means 4 discriminates the content of each item and an output selection means 5 selects an optimum transfer method corresponding to the importance and selects the content in matching with the method. Thus, the transfer of information by a facsimile equipment is secured and a job relating to the transfer is automated.

(19)日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11)特許出願公開番号

特開平7-273922

(43)公開日 平成7年(1995)10月20日

(51) Int.Cl.⁶
H 0 4 N 1/00
G 0 6 K 9/00

識別記号 104 Z
府内整理番号 Z 9289-5L

F I

技術表示箇所

(21)出願番号 特願平6-60762

(22)出願日 平成6年(1994)3月30日

審査請求 未請求 請求項の数4 O L (全6頁)

(71)出願人 000005821

松下電器産業株式会社
大阪府門真市大字門真1006番地

(72)発明者 伊藤 哲

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72)発明者 丹羽 寿男

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(72)発明者 萱嶋 一弘

大阪府門真市大字門真1006番地 松下電器
産業株式会社内

(74)代理人 介理士 小綱治 明 (外2名)

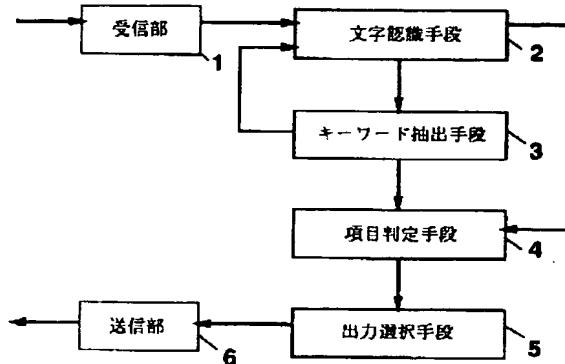
(54)【発明の名称】 FAX情報転送装置

(57)【要約】

【目的】 FAX送信表に自由に書かれた内容を項目毎にチェックして重要度に合わせた転送を行なう。

【構成】 受信部1に届いたFAXを文字認識手段2がテキストに変換し、キーワード抽出手段3がキーワードを抽出することによって文字認識率の向上と内容の要約を行なうとともに、項目判定手段4が項目毎の内容を判断し、出力選択手段5が重要度に応じた最適な転送方法とその方法に合わせた内容を選択する。

【効果】 FAXによる情報の伝達を確実にし、転送に関わる業務を自動化できる。



1

【特許請求の範囲】

【請求項1】FAX受信した画像情報を電子化された文字に変換する文字認識手段と、帳票上の各項目の文字列とその位置から各項目の内容を判定する項目判定手段と、項目毎の内容に基づいて受信画像および電子化された文字列の出力先と出力方法を決定する出力選択手段とを具備することを特徴とするFAX情報転送装置。

【請求項2】文字認識手段により認識された文字からキーワードを抽出し、このキーワードを転送することによって情報の圧縮を可能とするキーワード抽出手段を設けたことを特徴とする請求項1記載のFAX情報転送装置。

【請求項3】項目判定手段において、確定された項目内にその項目を判定するために使った単語以外の未知の単語が存在する場合、それを自動的に項目判定用の単語として登録する請求項1記載のFAX情報転送装置。

【請求項4】出力選択手段において、電子化された文字列の出力先として自由な文章が表示可能な携帯用情報端末を選択肢として持ち、これに電子化された文字列または抽出したキーワードを表示することを特徴とする請求項2記載のFAX情報転送装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】本発明は、FAXの受信および送信時に帳票の文字認識を行ない、自動的に受領したFAXの内容を転送する機能を持ったFAX情報転送装置に関する。

【0002】

【従来の技術】FAXで受信した画像をテキストに変換する従来のFAX-OCR装置は、扱う帳票のフォーマットが決められていて、所定の位置に特定の内容を書き込む必要がある。また、FAXで受信したり読み取った情報により自動的にFAXの送信や返信、転送、もしくは電話機やポケットベルによるFAX着信案内をする場合は、帳票に記載されたFAX番号や電話番号を読みとることで出力先を決めていた。

【0003】

【発明が解決しようとする課題】しかしながら上記従来の方法では、扱う帳票のフォーマットが限られている場合が多く、専用の記入用紙を必要とする上に、所定の位置に特定の内容を記入しなければ、情報を正しく入力することが出来ない。したがって、一般的のFAX通信において冒頭に送るFAX送信案内票などのように、自由なフォーマットで記入されているものを同様に扱うことはできなかった。

【0004】また、FAXで受信したり読み取った情報により自動的に送信や返信、転送などをする場合に、FAX番号や電話番号が記載されていない場合は、それを決定する手段がなく、また、送信先を自動的に決定できても、情報の内容や重要度によって出力方法を変

2

えるということが出来ない。

【0005】例えば、至急の用件のFAXや重要な案件のFAXが着信した場合は、電話で担当者を呼び出すことや出先に転送することが必要になるが、親元のFAXや特に重要でなく定期的に送られてくるようなFAXの着信時には、せいぜい本体の周辺にいる人だけにわかる程度で構ないので、情報の種類によらず単一の方法でしか情報を出力できない従来の方法では、業務を著しく阻害する恐れがあった。

【0006】本発明は、上記問題点を解決するために、自由なフォーマットのFAX送信案内票から認識すべき各項目を正しく認識し、そして各項目の情報の内容に合わせて出力先と出力手段を自動的に最適化するFAX情報出力装置を提供することを目的とするものである。

【0007】

【課題を解決するための手段】上記目的を達成するため、第1の発明は、受信したFAX画像を文字列に変換する文字認識手段と、帳票上の各項目の文字列と位置関係から各項目の内容を判定する項目判定手段と、項目毎の内容に基づいて、受信画像および電子化された文字列の出力先と出力方法を決定する出力選択手段とを備えた構成である。

【0008】第2の発明は、受信したFAX画像を文字列に変換する文字認識手段と、この文字列からキーワードを抽出するキーワード抽出手段と、帳票上の各項目の文字列と位置関係から各項目の内容を判定する項目判定手段と、項目毎の内容に基づいて、抽出されたキーワード情報の出力先と出力方法を決定する出力選択手段とを備えた構成である。

【0009】第3の発明は、受信したFAX画像を文字列に変換する文字認識手段と、この文字列からキーワードを抽出するキーワード抽出手段と、帳票上の各項目の文字列と位置関係から各項目の内容を判定する項目判定手段と、各項目に含まれるキーワードを自動的に項目判定手段において項目の属性を判定するための文字列として登録する手段と、項目毎の内容に基づき受信画像および電子化された文字列の出力先と出力方法を決定する出力選択手段とを備えた構成である。

【0010】第4の発明は、受信したFAX画像を文字列に変換する文字認識手段と、必要に応じて文字列からキーワードを抽出するキーワード抽出手段と、帳票上の各項目の文字列と位置関係から各項目の内容を判定する項目判定手段と、項目毎の内容に基づき受信画像および電子化された文字列の出力先と出力方法を決定し、文字列やキーワードの出力先として文字列が表示できる携帯用情報端末が選択可能な出力選択手段とを備えた構成である。

【0011】

【作用】上記した構成によって第1の発明では、FAX-OCR装置にFAXが着信し、電子的なテキストに変

3

換されると、帳票上の各項目の位置関係から項目判定手段が各項目の内容を判定し、各項目の内容に基づいて出力選択手段が出力先と出力方法を決定することによって、FAXの受信・送信時に自由なフォーマットの帳票に書かれた各項目から情報を収集することが可能で、利用範囲が広く、最適な出力先と出力方法の選択が可能なため、FAXを介した効率的な情報処理が実現できる。

【0012】第2の発明では、FAXが着信し、電子的なテキストに変換されると、キーワード抽出手段がキーワードを抽出し、このキーワードと帳票上の各項目の位置関係から項目判定手段が各項目の内容を判定し、各項目の内容に基づいて出力選択手段が出力先と出力方法を決定し、出力先の機器にキーワードを送ることによって、第1の発明の作用に加えて、キーワードを抽出することで情報の重要度の判定が容易となり、さらに出力する情報にキーワードを選択することで情報の圧縮も可能となる。

【0013】第3の発明では、項目判定手段において、項目判定に使われる単語以外の未知の単語を、次回以降の項目判定の単語として採用することにより、第1の発明の作用に加えて、自動的に項目判定の精度が上がる。

【0014】第4の発明では、出力選択手段において、電子化された文字列の出力先として文字列が表示可能な携帯用情報端末を出力先として選択できることにより、情報の確認を迅速にする効果が得られ、さらに抽出されたキーワードを表示することで、限られた表示スペースで全文を表示せずに内容が推定でき、転送する情報量を最小限に抑える。

【0015】

【実施例】以下、本発明の実施例について図面を参照しながら説明する。図1は本発明の一実施例におけるFAX情報出力装置を示すブロック図、図2は同実施例の文字認識手段のブロック図、図3は同実施例のキーワード抽出手段のブロック図、図4は同実施例の項目判定手段のブロック図、図5は同実施例の出力選択手段のブロック図である。

【0016】本実施例にかかるFAX情報転送装置は図1に示すとおり、受信部1において受信した画像情報を電子化された文字に変換する文字認識手段2と、文字認識された文字からキーワードを抽出するキーワード抽出手段3と、帳票上の各項目の文字列と位置から各項目の内容を判定する項目判定手段4と、得られた項目毎の内容に基づいて、これらの情報の出力先と出力方法を決定する出力選択手段5と、決定された出力方法で送信する送信部6とで構成されている。

【0017】ここに、文字認識手段2による処理の概要を図2により説明する。図2において、文字認識手段2は、文字切り出し部21と文字認識部22と後処理部23の各々の処理を順に行なうように構成されている。

【0018】文字切り出し部21は受信部1において展

10

20

30

40

50

4

開されたFAX画像の濃淡やその配置といった情報から一文字ずつ文字画像を切り分け、文字認識部22に出力するものである。このときFAX画像から得られる文字列の位置と罫線の情報は項目判定手段4に送られる。

【0019】次に文字認識手段22は、文字切り出し部21から送られてくる1文字ずつの文字画像をあらかじめ学習しておいた文字パターンと比較することにより、一文字につき複数の文字候補を後処理部23に出力する。

【0020】次に後処理部23では、一文字につき複数の文字候補をもつ文字列の全ての組合せを作り、それを辞書24に記載されている文字列と文法的な規則に照会することで正しい文字列の組合せを決定する。このとき、最も信頼度の高い文字列を正解候補とし、キーワード抽出手段3に結果を送る。

【0021】ここに、キーワード抽出手段3による処理の概要を図3により説明する。図3において、キーワード抽出手段3は、文字認識手段2から入力された文字情報を形態素解析部31が辞書33を用いてキーワード候補を作成し、統計処理部32が単語生起頻度ファイル34を参照してキーワードを抽出するように構成されている。

【0022】形態素解析部31は、文字情報を単語レベルまで解析し、キーワードとなり得る単語をキーワード候補として抽出する。このとき、辞書33に載っている単語の知識を参考する。辞書33の内容としてはキーワードとして選ばれることの多い品詞である名詞の文字表記が最低限必要である。なお、文字認識手段2の後処理部23における解析結果をそのまま使うか、辞書24を使うことにより、辞書33を用いた形態素解析の代替とすることもできる。

【0023】次に統計処理部32は、キーワード候補各々について帳票中の出現頻度Xと、単語生起頻度ファイル34の該当する単語生起頻度Yとからそのキーワード候補の尤度Eを計算し、尤度が一定の値を越えるか、もしくは尤度の多い順に一定数の候補だけキーワードとして出力する機能を持つ。単語生起頻度ファイル34は、単語がある文書内でどの位の頻度で出現する確率があるかを調べた統計値である単語生起頻度を記憶している。

【0024】キーワード尤度の算出例は $E = aX - bY$ (a, bは定数で、用途に合わせて決められる)等があり、こうすることで出現頻度Xのみでキーワードを決定する方式に比べ、どんな文書にも高頻度で出現するがキーワードとして相応しくないありふれた単語がキーワードとして選択されることを防止する効果がある。また、単語生起頻度は、その文書が属している分野毎に違いがあるので、用途によっては単語生起頻度ファイル34を分野別に用意し、分野に応じて選択しても良い。

【0025】また、キーワード抽出手段3で抽出される

5

キーワードをフィードバックし参照することによって、後処理部23での精度を向上することが可能である。これは、キーワードとして抽出されることの多い文字列は、そうでない文字列よりも正解である可能性が高いためである。

【0026】ここに、項目判定手段4による処理の概要を図4により説明する。図4において、項目判定手段4は、文字認識手段2において得られた文字列とその位置および罫線の情報より各項目を分離する項目分離部41と、それら分離された項目内の特定の語句より各項目の内容を推定する項目分類部42の処理を行なうように構成されている。

【0027】項目分離部41は、文字認識手段2において得られた文字列とその位置および罫線の情報より各項目を分離するように動作する。具体的には、文字列と文字列の間隔が空白により大きく開いていて、かつそれらの文字列のつながりがない所や、罫線によって区切られている所を境界線として、各々別の項目として扱うようになる。

【0028】次に、項目分類部42は、項目分離部41において分離された各々の項目に含まれる語句を属性キーワードファイル43に登録された語句に参照することによって、その項目にどんな内容が書かれているかを判定する。この属性キーワードファイル43には、取引先の会社名や担当者などの固有名詞の他にFAX送信票に特徴的に使われる各項目の内容を示すような用語（例えばFAXの宛先を示すものとして”送付先””受信者””受取人””To:”などが挙げられ、また、FAXの送り主を示す語句として”発信””発信元””発信者””発信人””From:””送信先””送付元”なども挙げられる）をそれぞれ登録しておく。

【0029】これらの属性キーワードはユーザーによる修正作業によって学習させることも可能であるが、判定された各項目にキーワード抽出手段3によって得られたキーワードが含まれている場合、このキーワードを属性キーワードファイル43に追加することで自動化が可能である。

【0030】ここに、出力選択手段5による処理の概要を図5により説明する。図5において、出力選択手段5は、項目判定手段4によって分類された各項目の属性について受信側情報ファイル52を参照することで、受信したFAX情報の転送先とその手段を決定する参照部51と、出力方法に合わせてFAX情報からどのような内容を出力するかを決定する情報選択部53とで構成されている。

【0031】参照部51は、項目判定手段4が分類した各項目の属性に基づいて、受信側情報ファイル52を参照する。受信側情報ファイル52には、受信者の氏名、所属、役職、電話番号、FAX番号の他に、送信元になり得る取引先や知人などの情報、重要度チェック用

50

6

キーワード、不在時の転送先、連絡方法などが記憶されている。特に転送先や連絡方法は、情報の重要度に合わせて受信者が一番好む転送先やその方法を複数登録しておく。重要度チェック用キーワードは、例えば、「至急」、「親展」、「見積り」など、そのFAXの内容を決めるものや、受信者が重要と判断した特定の会社名、製品名、氏名などの固有名詞などをあらかじめ重要度別に登録したものである。

【0032】まず、受信者の候補が複数いる場合、参照部51は受信側情報ファイル52の受信者の氏名、所属、役職、電話番号などを比較し受信者を特定することで転送先を決定する。そして、送信者などの各項目について受信側情報ファイル52に設定されている情報に対する重要度を参照して重要度を決定する。

【0033】次に、情報選択部53は、重要度に合わせて出力する情報の形態が異なるようにする。例えば、重要でないFAXは、FAX画像を記録紙に印刷し、FAX本体を発呼させる程度で良いが、重要であれば、受信者の携帯している情報端末に情報を転送するように、転送方法により転送する情報の内容を受信側情報ファイル52に記載された内容で決定する。

【0034】このようにして転送先と転送方法、転送内容が決定されたのち、電話回線をつないで情報を外部に送る送信部6に送られ、所定の場所に転送される。図6に示した出力先の例のように、携帯電話機やポケットベルを発呼させることによりFAXの着信をより確実に知らせるようにする。このとき、電話機に合成音声による着信案内をてもいいし、ポケットベルの自由文表示タイプのものであれば、そのままテキストに変換されたものを送ってもよいが、キーワード抽出手段3において得られたキーワードを使えば、より少ない字数で内容を認知させることが可能である。携帯情報端末として他には通信機能のある電子手帳なども考えられる。電子化された情報は記憶手段を用いることにより再利用可能である。

【0035】なお、本発明の各手段、各部は、コンピュータを用いてソフトウェア的に実現し、またはそれら各機能を有する専用のハード回路を用いて実現することが出来る。また、項目分離部41は、文字列とその位置および罫線の情報より各項目を分離する構成としたが、罫線のない帳票の場合は文字列とその位置より各項目を分離する構成としてもよい。

【0036】

【発明の効果】以上述べたことから明らかなように、本発明は、FAXの受信者側がそのFAXの内容を意識することなくもっとも理想的な方法によりFAX転送や着信案内を受けることが可能となる。

【0037】また、文字認識手段にキーワード抽出結果をフィードバックした場合は、画像から文字列に変換する際に多少の誤認識があっても、比較的の信頼性の高い情

7

報として扱うことが可能となるばかりでなく、キーワードを転送する情報として使うことで情報の圧縮が可能となる。

【図面の簡単な説明】

【図1】本発明の一実施例におけるFAX情報転送装置を示すブロック図

【図2】同実施例の文字認識手段の処理を説明するためのブロック図

【図3】同実施例のキーワード抽出手段の処理を説明するためのブロック図

【図4】同実施例の項目判定手段の処理を説明するためのブロック図

8

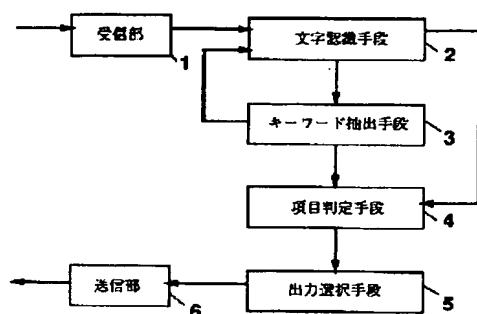
【図5】同実施例の出力選択手段の処理を説明するためのブロック図

【図6】同実施例の送信部の出力先の例を説明するためのブロック図

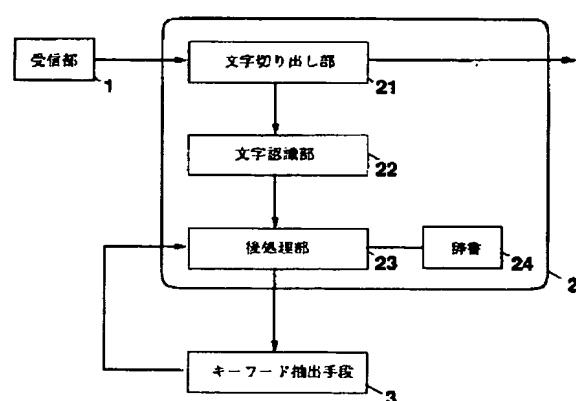
【符号の説明】

- 1 受信部
- 2 文字認識手段
- 3 キーワード抽出手段
- 4 項目判定手段
- 5 出力選択手段
- 6 送信部

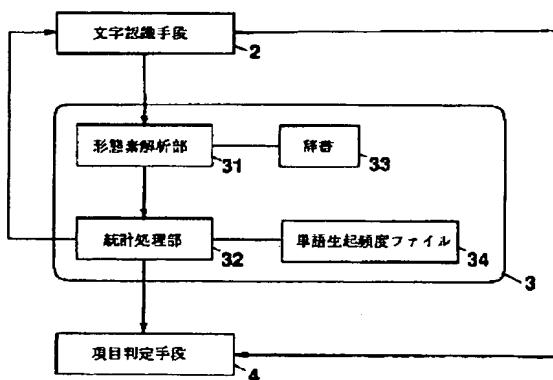
【図1】



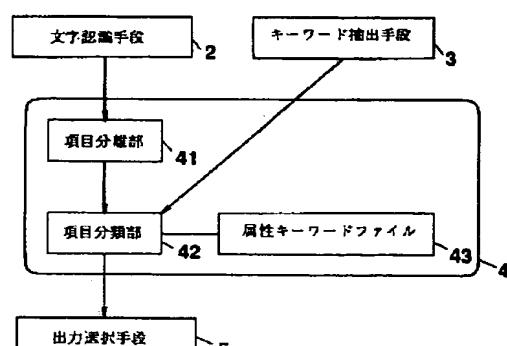
【図2】



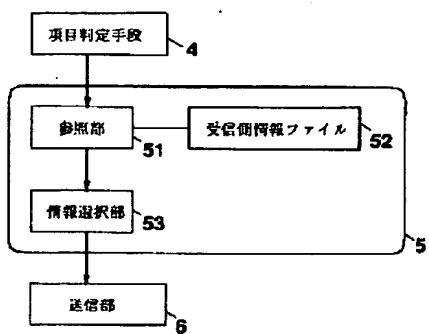
【図3】



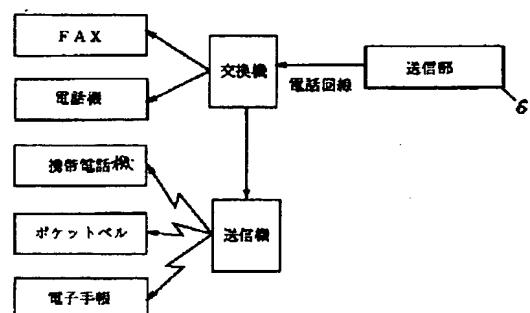
【図4】



【図5】



【図6】



[Claim(s)]

[Claim 1] FAX information-transfer equipment characterized by to provide an item judging means judge the contents of each item from a character recognition means to change into the alphabetic character which had the image information which carried out FAX reception electronized, character string, and location of each item on a document, and an output selection means determine a receiving image, and the output destination change and output method of the electronized character string based on the contents for every item.

[Claim 2] FAX information transfer equipment according to claim 1 characterized by establishing the keyword extraction means which enables informational compression by extracting a keyword from the alphabetic character recognized by the character recognition means, and transmitting this keyword.

[Claim 3] FAX information transfer equipment according to claim 1 which registers it as a word for an item judging automatically when strange words other than the word used in order to judge the item in the fixed item in an item judging means exist.

[Claim 4] FAX information transfer equipment according to claim 2 characterized by displaying the character string which had the portable information terminal which can display a text free as an output destination change of the electronized character string in an output selection means as a selection branch, and was electronized by this, or the extracted keyword.

[Detailed Description of the Invention]

[0001]

[Industrial Application]

This invention performs character recognition of a document at the time of reception of FAX, and transmission, and relates to FAX information transfer equipment with the function to transmit the contents of FAX received automatically.

[0002]

[Description of the Prior Art]

It opts for the format of the document to treat and the conventional FAX-OCR equipment which changes into a text the image received by FAX needs to write the specific contents in a position. Moreover, when FAX arrival guidance by transmission, a reply, a transfer or telephone, and pocket bell of FAX was automatically carried out using the information which received by FAX or was read, the output destination change was decided by reading the FAX number indicated by the document and the telephone number.

[0003]

[Problem(s) to be Solved by the Invention]

However, by the above-mentioned conventional approach, the format of the document to treat is restricted in many cases, and if the entry form of dedication is needed upwards and the specific

contents are not written down in a position, information cannot be inputted correctly. Therefore, similarly what is filled in in a free format was not able to be treated like the FAX transmitting guidance vote sent to the beginning in a general FAX communication link.

[0004] Moreover, even if there is no means to determine it and it is able to determine a transmission place automatically when carrying out transmission, a reply, a transfer, etc. automatically using the information which received by FAX or was read, and neither a FAX number nor the telephone number is indicated, it cannot be said that an output method is changed with the informational contents and significance.

[0005] For example, although it is necessary to transmit to calling a person in charge by telephone or a destination when FAX of urgent business and FAX of an important issue receive a message At the time of the arrival of confidential FAX or FAX which is sent especially importantly and periodically, since extent which only those who are around a body at most make known was available, there was a possibility of checking business remarkably, by the conventional approach which is not based on an informational class but can output information only by the single approach.

[0006] This invention aims at offering the FAX information output unit which recognizes correctly each item which should be recognized from the FAX sending guidance vote of a free format, and optimizes an output destination change and an output means automatically according to the contents of the information on each item, in order to solve the above-mentioned trouble.

[0007]

Means for Solving the Problem

In order to attain the above-mentioned purpose, the 1st invention is the configuration equipped with an item judging means to judge the contents of each item, and an output selection means to determine a receiving image, and the output destination change and output method of the electronized character string based on the contents for every item, from a character recognition means to change the received FAX image into a character string, and the character string and physical relationship of each item on a document.

[0008] The 2nd invention is the configuration equipped with an item judging means to judge the contents of each item from a character recognition means to change the received FAX image into a character string, a keyword extraction means to extract a keyword from this character string, and the character string and physical relationship of each item on a document, and an output selection means to determine the output destination change and output method of the keyword information extracted based on the contents for every item.

[0009] A character recognition means to change into a character string the FAX image which the 3rd invention received, An item judging means to judge the contents of each item from the character string and physical relationship of each item on a document to be a keyword extraction means to extract a keyword from this character string, It is the configuration equipped with the

output destination change of a receiving image and the electronized character string, and an output destination change selection means to determine an output method, based on a means to register the keyword contained in each item as a character string for judging the attribute of an item in an item judging means automatically, and the contents for every item.

[0010] A character recognition means to change into a character string the FAX image which the 4th invention received, A keyword extraction means to extract a keyword from a character string if needed, An item judging means to judge the contents of each item from the character string and physical relationship of each item on a document, The portable information terminal which determines a receiving image, and the output destination change and output method of a character string which were electronized based on the contents for every item, and can display a character string as an output destination change of a character string or a keyword is the configuration equipped with the selectable output selection means.

[0011]

[Function]

If FAX receives a message to FAX-OCR equipment and it is changed into an electronic text by the 1st invention by the above-mentioned configuration When an item judging means judges the contents of each item from the physical relationship of each item on a document and an output selection means determines an output destination change and an output method based on the contents of each item It is possible to collect information from each item written to the document of a format free at the time of reception and transmission of FAX, the use range is wide, and since selection of the optimal output destination change and an output method is possible, the efficient information processing through FAX is realizable.

[0012] If FAX receives a message and it is changed into an electronic text in the 2nd invention A keyword extraction means extracts a keyword and an item judging means judges the contents of each item from the physical relationship of each item on this keyword and a document. By an output selection means' determining an output destination change and an output method based on the contents of each item, and sending a keyword to the device of an output destination change In addition to an operation of the 1st invention, the judgment of an informational significance becomes easy by extracting a keyword, and informational compression also becomes possible by choosing a keyword as the information outputted further.

[0013] In addition to an operation of the 1st invention, in an item judging means, the precision of an item judging goes up by 3rd invention automatically by adopting strange words other than the word used for an item judging as a word of a subsequent section eye judging next time.

[0014] The effectiveness which makes an informational check quick by the ability being able to choose the electronized portable information terminal which can display a character string as an output destination change of a character string as an output destination change in an output selection means in the 4th invention is acquired, the contents can be presumed by displaying the keyword extracted further, without displaying a whole sentence in the limited display tooth

space, and the amount of information to transmit is suppressed to the minimum.

[0015]

Example

Hereafter, it explains, referring to a drawing about the example of this invention. For the block diagram and drawing 2 which show a FAX information output unit [in / in drawing 1 / one example of this invention], the block diagram of the character recognition means of this example and drawing 3 are [the block diagram of the item judging means of this example and drawing 5 of the block diagram of the keyword extraction means of this example and drawing 4] the block diagrams of the output selection means of this example.

[0016] A character recognition means 2 to change into the alphabetic character which had the image information which received in the receive section 1 electronized as the FAX information transfer equipment concerning this example was shown in drawing 1 , A keyword extraction means 3 to extract a keyword from the alphabetic character by which character recognition was carried out, It consists of the transmitting sections 6 which transmit from the character string and location of each item on a document with the output method determined as an item judging means 4 to judge the contents of each item, and an output selection means 5 to determine the output destination change and output method of such information based on the contents for every obtained item.

[0017] Drawing 2 explains the outline of processing by the character recognition means 2 here. In drawing 2 , the character recognition means 2 is constituted so that each processing of the alphabetic character logging section 21, the character recognition section 22, and the after-treatment section 23 may be performed in order.

[0018] The alphabetic character logging section 21 carves a single-character [every] alphabetic character image from the shade of the FAX image developed in the receive section 1, or the information of the arrangement, and outputs it to the character recognition section 22. The location of a character string and the information on a ruled line which are acquired from a FAX image at this time are sent to the item judging means 4.

[0019] Next, the character recognition means 22 outputs two or more alphabetic character candidates to the after-treatment section 23 per single character by comparing with the character pattern which learned beforehand the alphabetic character image per character sent from the alphabetic character logging section 21.

[0020] Next, all the combination of the character string which has two or more alphabetic character candidates per single character is made from the after-treatment section 23, and the combination of a right character string is determined by referring for it to the character string indicated by the dictionary 24 and a grammatical regulation. At this time, a character string with the highest reliability is made into a correct answer candidate, and a result is sent to the keyword extraction means 3.

[0021] Drawing 3 explains the outline of processing by the keyword extraction means 3 here. In

drawing 3 , the morphological analysis section 31 creates a keyword candidate for the text inputted from the character recognition means 2 using a dictionary 33, and the keyword extraction means 3 is constituted so that the statistics processing section 32 may extract a keyword with reference to the word occurrence frequency file 34.

[0022] The morphological analysis section 31 analyzes text to word level, and extracts the word which can serve as a keyword as a keyword candidate. At this time, the knowledge of the word which appears in the dictionary 33 is referred to. The alphabetic character notation of the noun which is a part of speech with being chosen [much] as a keyword as contents of the dictionary 33 is the need at worst. In addition, it can also consider as an alternative of the morphological analysis using a dictionary 33 by using a dictionary 24, using the analysis result in the after-treatment section 23 of the character recognition means 2 as it is.

[0023] Next, the statistics processing section 32 calculates the keyword candidate's likelihood E about the keyword candidates of each from the frequency of occurrence X in the inside of a document, and the word occurrence frequency Y where the word occurrence frequency file 34 corresponds, and has the function which likelihood exceeds a fixed value or only a fixed number of candidates output to order with much likelihood as a keyword. The word occurrence frequency file 34 has memorized the word occurrence frequency which is the statistic which investigated whether there would be any probability to appear by the frequency of which about within a document with a word.

[0024] The example of calculation of keyword likelihood has $E=aX-bY$ (a and b are constants and are decided according to an application) etc., and compared with the method which determines a keyword with the frequency of occurrence X by carrying out like this, although it appears in any documents by high frequency, it is effective in preventing that the common word which is not suitable as a keyword is chosen as a keyword. Moreover, since word occurrence frequency has a difference for every field to which the document belongs, it may prepare the word occurrence frequency file 34 according to a field depending on an application, and may choose it according to a field.

[0025] Moreover, it is possible to improve the precision in the after-treatment section 23 by referring to the keyword extracted with the keyword extraction means 3 fee DOBBAKKU. The character string with this being extracted [much] as a keyword is because possibility of being a correct answer is higher than the character string which is not so.

[0026] Drawing 4 explains the outline of processing by the item judging means 4 here. In drawing 4 , the item judging means 4 is constituted so that the item separation section 41 which separates each item from the information on the character string obtained in the character recognition means 2, its location, and a ruled line, and the item classification section 42 which presumes the contents of each item from the specific phrase within the they-separated item may be processed.

[0027] The item separation section 41 operates so that each item may be separated from the information on the character string obtained in the character recognition means 2, its location,

and a ruled line. Spacing of a character string and a character string is greatly open with the null, and it is made to specifically treat as another item respectively by making into a boundary line a place without relation of those character strings, and the place divided by the ruled line.

[0028] Next, the item classification section 42 judges what kind of contents are written to the item by referring to the phrase contained in each item separated in the item separation section 41 in the phrase registered into the attribute keyword file 43. the vocabulary ("To:" etc. is mentioned as what shows the destination of FAX -- " -- a receiver's address ""addressee"" recipient --) as shows the contents of each item used for this attribute keyword file 43 characteristic of the FAX transmitting vote other than proper nouns, such as a customer's firm name and a person in charge moreover -- as the phrase which shows the sender of FAX -- "dispatch" -- "a sending agency ""addresser"" addresser" -- "From: ""transmission place"" sending origin" etc. is mentioned -- having -- it registers, respectively.

[0029] Although it is possible to also make it learn by the correction by the user, these attribute keywords can be automated by adding this keyword to the attribute keyword file 43, when the keyword obtained by the keyword extraction means 3 is contained in each judged item.

[0030] Drawing 5 explains the outline of processing by the output selection means 5 here. In drawing 5 , the output selection means 5 is referring to the receiving-side information file 52 about the attribute of each item classified according to the item judging means 4, and consists of the reference section 51 which determines the destination and the means of the received FAX information, and the information selection section 53 which determines what kind of contents are outputted from FAX information according to an output method.

[0031] Refer to the receiving-side information file 52 for the reference section 51 based on the attribute of each item which the item judging means 4 classified. The information of the customer who can become a transmitting agency else [, such as an addressee's name, affiliation, an executive, the telephone number and a FAX number,], an acquaintance, etc., the keyword for a significance check, the destination at the time of an absence, the communication approach, etc. are memorized by the receiving-side information file 52. The destination and the communication approach register especially two or more destinations which an addressee likes most, and its approaches according to an informational significance. Proper nouns, such as a specific firm name judged that what determines the contents of the FAX, such as "confidential" and a "estimate", and an addressee are important for the keyword for a significance check "urgently", for example, a product name, and a name, etc. are beforehand registered according to significance.

[0032] First, when there are two or more candidates of an addressee, the reference section 51 determines the destination by comparing the name of the addressee of the receiving-side information file 52, affiliation, an executive, the telephone number, etc., and specifying an addressee. And with reference to the significance over the information set as the receiving-side information file 52, significance is determined about each item, such as a transmitting person.

[0033] Next, it is made for the gestalten of the information which outputs the information

selection section 53 according to significance to differ. For example, unimportant FAX prints a FAX image on the recording paper, and although it is good at extent to which call origination of the body of FAX is carried out, if it is important, it will determine the contents of the information transmitted by the transfer approach from the contents indicated by the receiving-side information file 52 so that information may be transmitted to the information terminal which the addressee is carrying.

[0034] Thus, after the destination, the transfer approach, and the contents of a transfer are determined, it is sent to the transmitting section 6 which connects the telephone line and sends information outside, and is transmitted to a predetermined location. Like the example of the output destination change shown in drawing 6 , the arrival of FAX is more certainly told by carrying out call origination of a portable telephone or the pocket bell. At this time, arrival-of-the-mail guidance by synthesized speech may be carried out to telephone, as long as it is the free sentence display type thing of a pocket bell, what was changed into the text as it was may be sent, but if the keyword obtained in the keyword extraction means 3 is used, it is possible to make the contents recognize with smaller number of letters. The electronic notebook which has communication facility in others as a Personal Digital Assistant is considered. The electronized information is reusable by using a storage means.

[0035] In addition, each means of this invention and each part are realizable using the hard circuit of the dedication which is realized by software using a computer or has each [these] function. Moreover, although the item separation section 41 considered as the configuration which separates each item from the information on a character string, its location, and a ruled line, in the case of a document without a ruled line, it is good also as a configuration which separates each item from a character string and its location.

[0036]

[Effect of the Invention]

This invention becomes possible [receiving a FAX transfer and arrival-of-the-mail guidance by the most ideal approach], without the addressee side of FAX being conscious of the contents of the FAX so that clearly from having stated above.

[0037] Moreover, when a keyword extraction result is fed back to a character recognition means, in case it changes into a character string from an image, even if there is some incorrect recognition, it becomes compressible [information] by it not only becoming possible to treat as comparatively reliable information, but using as information which transmits a keyword.

[Brief Description of the Drawings]

[Drawing 1] The block diagram showing the FAX information transfer equipment in one example of this invention

[Drawing 2] The block diagram for explaining processing of the character recognition means of this example

[Drawing 3] The block diagram for explaining processing of the keyword extraction means of this example

[Drawing 4] The block diagram for explaining processing of the item judging means of this example

[Drawing 5] The block diagram for explaining processing of the output selection means of this example

[Drawing 6] The block diagram for explaining the example of the output destination change of the transmitting section of this example

[Description of Notations]

1 Receive Section

2 Character Recognition Means

3 Keyword Extraction Means

4 Item Judging Means

5 Output Selection Means

6 Transmitting Section